IMPACT OF TASK LOAD AND GAZE ON SITUATION AWARENESS IN UNMANNED AERIAL VEHICLE CONTROL

Joseph T. Coyne
Ciara M. Sibley
Naval Research Laboratory
Washington, DC

Increasing levels of automation and rising costs of manpower are pushing the DoD towards a supervisory control paradigm for future unmanned aerial vehicle (UAV) missions. Using the Supervisory Control Operations User Testbed, a group of 20 participants completed two twenty minute supervisory control missions where eye tracking and performance data were collected. Each mission had 3 levels of task load; which were manipulated by varying the frequency of events to which the user responded. During each level, the simulation paused and a situation awareness (SA) probe appeared with all UAVs and targets randomly placed on the map. Participants were tasked to reconstruct the map. Results showed higher load was associated with a significant decrease in SA. Additionally, participants spent significantly less time looking at the map when the task load was high. These results suggest that eye gaze may be a useful predictor of SA within a supervisory control task.

Unmanned Aerial Vehicles (UAVs) accounted for only 5% of the Department of Defense's (DoD) aircraft inventory in 2005, however by 2012 that number has surged to 41% (Gertler, 2012). The rapid rise in vehicles and the simultaneous reduction in force within the DoD has increased the interest in shifting the UAV operations paradigm away from direct control of specific vehicle functions (e.g., piloting and payload) and towards a supervisory control model where a single operator monitors multiple vehicles (DoD, 2013). Increases in UAV automation, in addition to open system architectures will help drive this change. UAV control has already begun to shift from control via "stick and rudder" to waypoint navigation. The majority of an operator's tasking in this future environment will likely be to assign vehicles to different targets and objectives, monitor the progress of those vehicles, and update plans given new opportunities and changing information.

Problems associated with increased automation and decreased situation awareness (SA) have been discussed in a number of studies (e.g., Calhoun et al., Endsley & Kaber, 1999). Reduced SA primarily results from a combination of complacency and lack of interaction with the system, and these SA lapses can lead to an inability to detect problems when automation fails, as well as increased time to recover after an error. For example, Calhoun et al. (2011) investigated varying levels of automation and reliability within a UAV supervisory control environment and found that after several sessions in which automation performed perfectly, every participant missed a route error introduced by the automation, despite being told that the automation could make mistakes. While the importance of SA within UAV supervisory control has often been suggested to be a potential issue, the assessment of SA has been limited. Calhoun et al (2001) inferred poor SA based upon user's lack of a correct response, this represents an implicit measure of SA and is only one of several types of SA measures (Sarter & Woods, 1995). The focus of this paper is to investigate two other methods of assessing SA within a UAV supervisory control environment, specifically through SA probes and physiological measurement.

SA probes such as SAGAT (Endsley, 1998) are one of the most common methods of SA assessment. In these probes the environment is paused and hidden and the individual is asked a specific series of questions regarding its current state. While probe measures provide a direct means of measuring SA they can be disruptive, only provide data at discrete points in time, and can only be used within controlled studies. The ability to have continuous measures of SA, which also provide insight into the process behind acquiring and maintain SA are highly desirable. The use of physiological data,

specifically eye tracking based metrics, are beginning to gain traction as a continuous noninvasive SA measurement technique (e.g., Moore & Gugerty, 2010; Van de Merwe et al., 2012). While Van de Merwe couples gaze information with an implicit measures of SA (i.e., detection of an aircraft failure), Moore and Gugerty's research was one of the only studies to include both direct measures of SA (using SAGAT) and eye tracking. Mooore and Gugerty found that the percentage of time looking within areas of interest (aircraft within an ATC display) accounted for a significant portion of the variance within the SA probe, suggesting that eye tracking data is an effective means of assessing SA implicitly.

The purpose of the present study is to further investigate the utility of eye tracking metrics within a supervisory control environment. Specifically the authors hope to replicate the findings of previous researchers and demonstrate a relationship between eye tracking data and performance within a SA probe.

Method

Supervisory Control Operations User Testbed

The Supervisory Control Operations User Testbed (SCOUT) was developed by the Naval Research Laboratory to replicate the tasks that a UAV mission commander and air vehicle pilot will perform in future operations with increased automation. SCOUT was designed as a two screen game environment that enables a single operator to control 3 heterogeneous UAVs. The vehicles differ in their speed and sensor range (which influences the time to complete a target search). The tasks within SCOUT include navigation and route planning; airspace management (requesting access to controlled airspace); communication (responding to requests for information); and adjusting air vehicle parameters (i.e., altitude and speed) and target parameters (i.e., location and search radius) as new information and commands are issued. The route planning task was the participant's main priority and main source of points within the game. The task required the participant to assign multiple targets of varying values, deadlines, and uncertainty (potential time required to locate the target) to the three UAVs they were supervising. Each vehicle would automatically search for their assigned target once they arrived within the target area. The payload task was entirely automated where the target would be found and the corresponding points awarded once the vehicle was within sensor range of the target. In addition, the participant received points for responding to requests for information and vehicle commands. Points were reduced when the participant's vehicles entered into restricted airspace without requesting access. Figure 1 depicts SCOUT's main mission management screen. During the experiment, the map was locked in position and represented an area of approximately 65 x 45 kilometers.

Equipment

A SmartEye Pro 6.1 five camera system was used to capture eye tracking data from participants at a frequency of 60Hz. The data from the SmartEye system was sent via network packets to the computer running SCOUT and was integrated with the participants' behavioral and simulation data.

Experimental Design

Twenty civilian employees and summer interns at the Naval Research Laboratory volunteered for participation in the experiment. Participants completed a 30 minute SCOUT training session consisting of a series of videos and sample tasks. Upon completing the training, there were two experimental sessions. Each experimental session began with a planning phase in which participants had up to ten minutes to select their initial vehicle and target assignments before the vehicles began moving. After the planning phase, the participant completed an 18 minute experimental block which consisted of a six minute easy, medium, and hard segment, in which events were presented approximately every 75,45, and 15 seconds respectively. A situation awareness probe (depicted in Figure 2) was presented within each 6 minute block. During the SA probe the SCOUT control displays would disappear and the probe would

appear with the vehicles and targets randomly placed on the map. The participant had to two minutes to move the vehicles and targets to their estimated position on the map and then indicate which target each vehicle was currently pursuing and its value . Once participants completed the probe, the simulation would return and vehicles would only begin moving again once a “resume mission” button was pressed.

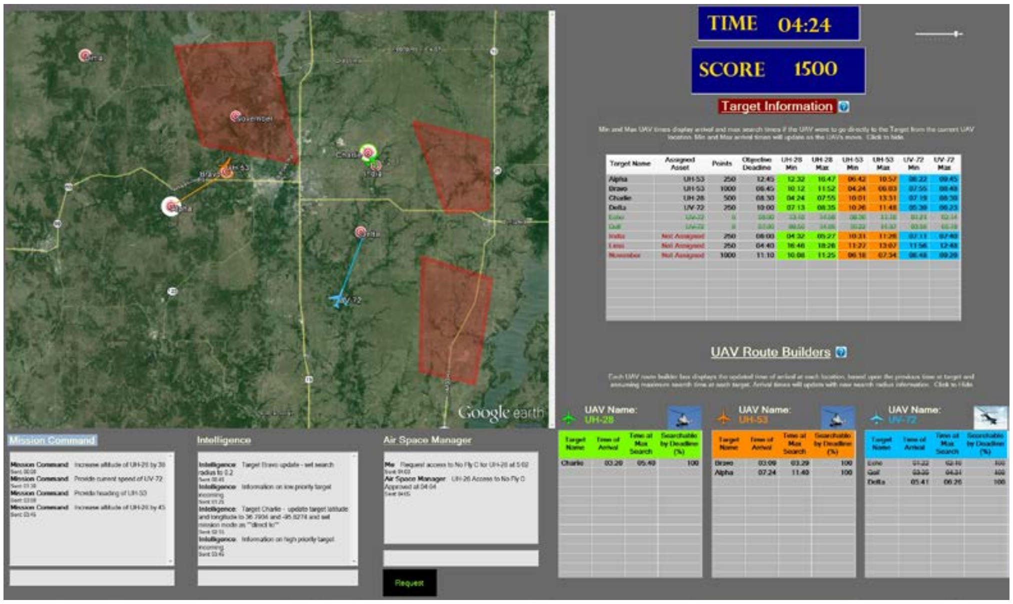

Figure 1. SCOUT’s mission management screen where operators assigned targets to their different vehicles.

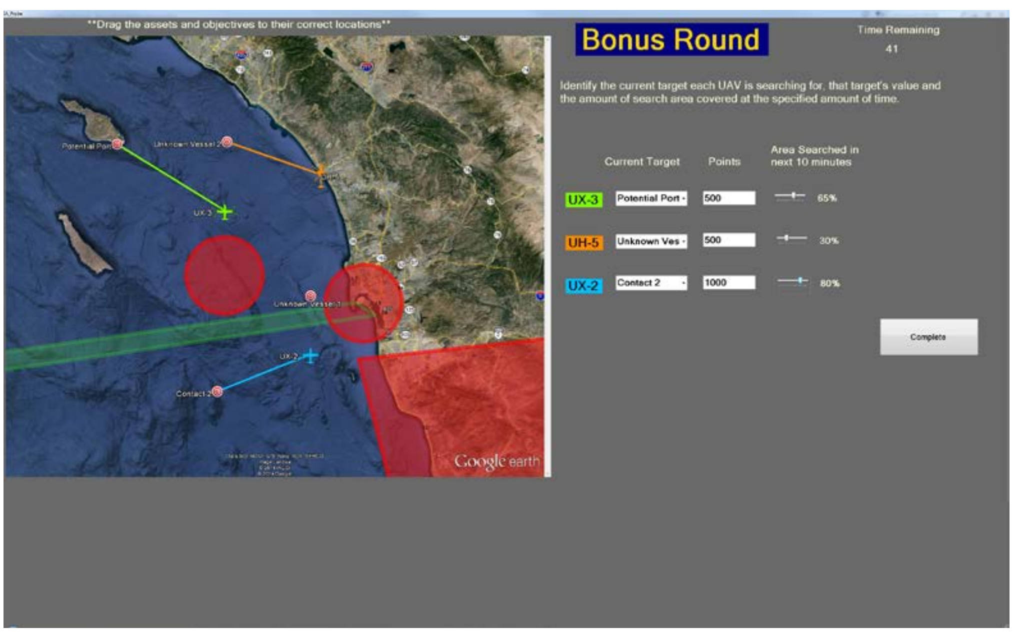

Figure 2. SCOUT’s situation awareness probe.

Results

Situation Awareness Probe Results

The primary metric from the SA probe was the distance between the participant's placement of the vehicles and targets on the map and the actual position of those objects in the simulation, immediately preceding the probe. The maximum possible error for each object was limited to 35 km. A two-way (session x difficulty) repeated measures ANOVA was run on the SA probe error data. There was a significant main effect of difficulty $F(2,38) = 3.382$, $p = .044$ (see Figure 3). A Tukey HSD post hoc analysis revealed that the SA probe error was significantly smaller in the easy task load compared to the difficult task load. There was no effect of session ($F(1,19) = 2.109$, $p = .163$) or interaction of session and difficulty on SA probe error ($F(2,38) = .834$, $p = .442$). An additional metric from the SA probe was the ability to correctly identify which target each vehicle was pursuing. The results of this analysis mirrored those of the error distance with only a main effect of difficulty being present $F(2,38) = 5.729$, $p = .007$. Performance within the east condition (61%) was significantly better than the hard condition (36%).

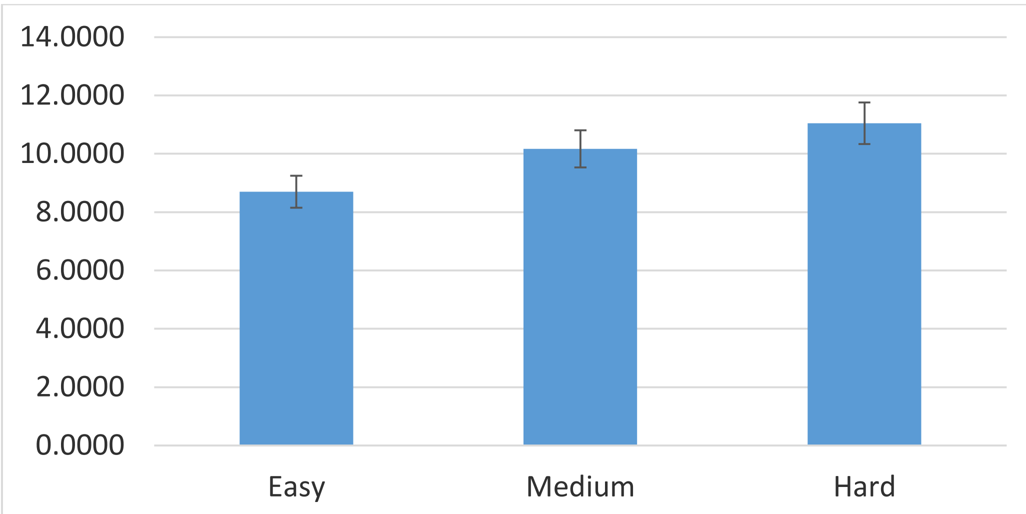

Figure 3. SA Probe Error distance across the three levels of task difficulty. Error bars represent the standard error of the mean.

Gaze Results

The eye tracking analysis focused on the data collected one minute prior to each SA probe. Specifically, the principal eye tracking metric was percentage of time spent looking at the map prior to the probe. A two-way (session x difficulty) repeated measures ANOVA was run on the percentage of map dwell time data. There was a significant main effect of difficulty $F(2,38) = 5.174$, $p = .010$ (see Figure 4). A Tukey HSD post hoc analysis revealed that participants spent significantly less time looking at the map in the hard condition compared to both the easy and medium conditions. There was no effect of session ($F(1,19) = 1.693$, $p = .209$) or interaction of session and difficulty ($F(2,38) = .755$, $p = .478$) on percentage of map dwell time.

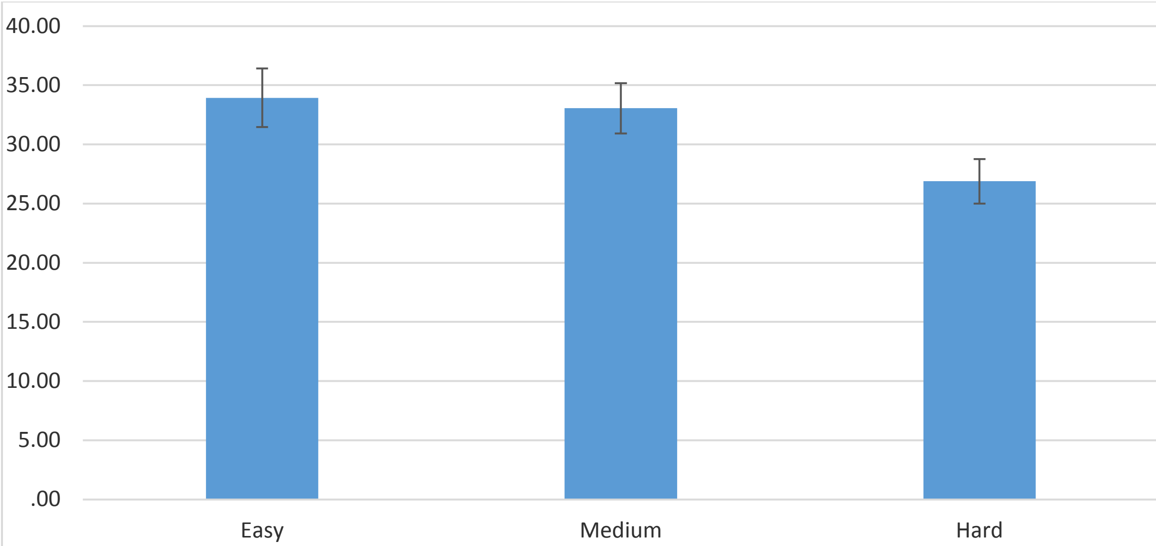

Figure 4. Percent of time spent looking at the map 1 minute prior to probe. Error bars represent the standard error of the mean.

Variable Correlations

Correlations were run to understand the relationships between the different variables of interest. In addition to Difficulty level, SA Error and Time in Map, the average amount of eye movement over the 1 minute time period in pixels was investigated as an additional measure (see Table 1).

Table 1.
Correlation matrix for SA Error, eye metrics and difficulty

	SA Error	Time in Map	Eye Movement	Difficulty
SA Error	1			
Time in Map	-.229*	1		
Eye Movement	.112	-.212*	1	
Difficulty	.215*	-.209*	.110	1

Note. * Indicates significance at .05 level

Discussion

The results of this study indicate that both performance on an SA probe and gaze were significantly impacted by task demands. As task demands increased, operators spent significantly less time looking at the map display (which aided in their ability to avoid restricted airspace and identify new targets of opportunity) and showed significant reductions in their SA. Further, these results support those of previous researchers (Moore & Gugerty, 2010) and provide evidence that eye tracking measures can be used as a potential supplement to direct measures of SA. While direct measures of SA offer high face validity, they are disruptive and difficult to implement in many environments. As such, identification and use of a non-obtrusive continuous measure of SA is of great value. Such a measure could help drive adaptive automation, such as tailored alerting (Ratwani, Mccurry & Trafton, 2010) or be useful in assessing operator performance with new displays or automation. Although the correlations between SA probe error and Time in Map were significant, they were small, suggesting that eye tracking alone cannot account for an individual's SA probe values and that such measures are still valuable in understanding SA.

It is important to note that within the present study, task demands were driven by chat communication as well as the inclusion of new targets of opportunities. These tasks often took the operator's attention away from the map display. The eye movement correlation data shows that as time spent looking at the map increased the amount of overall movement tended to decrease. It is unclear how queries or tasking that pushed attention to the map display would have impacted either the eye tracking metrics or the performance on the SA probe. Additionally the eye tracking analysis described here treated the entire map as a single area of interest. Future research may look at time on each object on the map and SA probe performance for that specific object. The version of SCOUT used in this experiment made such analysis challenging, however improvements to the map functionality now allow for easy conversion of pixel location to both map objects and latitude and longitude. This new functionality will aid in a more precise measure of eye tracking and SA.

Requesting that participants reposition the icons on the map of the same dimension and size represents the most basic SA, i.e. level 1 perception of elements (Endsley, 1995). Future experiments within SCOUT will look at higher levels of SA. Potential changes such as repositioning the area within the map so that user can no longer rely on screen position but must interpret location may can help investigate level 2 SA (Comprehension), additionally requesting individuals draw the position of their aircraft at a future time can help assess level 3 SA (Projection). The use of position based SA probes provides an advantage over traditional SA queries within SAGAT (e.g., where was the aircraft landing) in that the measure is continuous.

Overall the present study helped confirm the utility of using eye tracking as a supplemental measure of operator state and awareness.

References

Calhoun, G.L., Draper, M.H., and Ruff, H.A. (2009). "Effect of level of automation on unmanned aerial vehicle routing task", in: *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp. 197-201) San Antonio, Texas.

Department of Defense (2013). "Unmanned Systems Integrated Roadmap FY2013-2038". (Washington, DC: Department of Defense).

Endsley, M.R. (1988). "Situation awareness global assessment technique (SAGAT)", in: *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference* (Dayton, OH: *IEEE*), 789-795.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors,* 37(1)**,** 32-64.

Endsley, M.R., and Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics* 42**,** 462-492.

Gertler, J. (2012). "U.S. Unmanned Aerial Systems". (Washington, DC: Congressional Research Service).

Moore, K. and Gugerty, L. (2010) Development of a novel measure of situation awareness: The case for eye movement analysis, in *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 1650-1654) San Francisco, CA.

Ratwani, R.M., Mccurry, J.M., and Trafton, J.G. (2010). "Single Operator, Multiple Robots: An Eye Movement Based Theoretic Model of Operator Situation Awareness", in: *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction* (pp. 235-242) Osaka, Japan.

Sarter, N. B. and Woods, D. D. (1995) How in the world did we get into that mode? Mode error and awareness in supervisory control. *Human Factors*, 37(1), 5-19.

Van De Merwe, K., Van Dijk, H., and Zon, R. (2012). Eye movements as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology* 22**,** 78-95.